

# Capítulo 8.

## Precisión de los Datos

---

### **EL ARCHIVO MAESTRO DE DIRECCIONES Y LOS PROCEDIMIENTOS DE ENUMERACIÓN**

El método de entrega de cuestionarios utilizado en Puerto Rico fue actualización/entrega. El listado de direcciones para las áreas de actualización/entrega se compiló durante una operación de campo del Negociado del Censo que se conoce como Listado de Direcciones. Los empleados del censo se mandaron al campo con mapas de su área de asignación y fueron instruidos a anotar las direcciones de tipo urbano, las direcciones de tipo no urbano o una descripción de la localización, o posiblemente alguna combinación de éstas, para cada unidad de vivienda. Además, la localización de la unidad se anotó en el mapa censal con lo que se conoce como un punto en el mapa. Esta operación se llevó a cabo en el otoño de 1998.

Al completar el procesamiento de datos sobre el listado de direcciones, fue posible tabular el número de unidades de vivienda en cada bloque. Como las unidades de vivienda en esas áreas podrían tener direcciones postales que no fueran estándares y que podrían ser anotadas en archivos censales sólo con una descripción de la localización, se le envió listas de recuentos de unidades de vivienda por bloque a las unidades gubernamentales que estaban participando en la operación de revisión local en esas áreas. Esta operación se conoció como LUCA 99. Cuando un participante de LUCA 99 no estaba de acuerdo con un recuento de bloques censales, se hizo un recorrido de nuevo de este bloque, durante el cual los empleados del censo fueron reasignados para actualizar el listado de direcciones. También, hubo un proceso de apelación para resolver discrepancias en la clasificación de unidades de vivienda, que podría tener como resultado añadir unidades de vivienda al listado de direcciones. El segundo recorrido de LUCA 99 y el proceso de apelación de LUCA 99 se llevaron a cabo varias veces durante la creación del Archivo Maestro Decenal de Direcciones. La mayoría de las entidades de LUCA 99 tenían procesados sus resultados del segundo recorrido antes de crear el Archivo Maestro Decenal de Direcciones inicial, pero muchos no lo tenían. Se diseñaron muchas actualizaciones al Archivo Maestro Decenal de Direcciones específicamente para añadirle resultados que se recibieron después del segundo recorrido y la apelación a los archivos censales para que se entregaran los cuestionarios a tiempo.

La última operación de hacer listados de direcciones en las áreas de actualización/entrega fue la operación Actualización/Entrega. Durante esta operación, se entregó a mano en cada unidad de vivienda un cuestionario del censo. Durante el proceso, se actualizaron el Archivo Maestro de Direcciones y los mapas.

Las personas que no recibieron un cuestionario en su residencia pudieron entregar un cuestionario de Sea Usted Contado o llamar a Asistencia por Teléfono sobre Cuestionarios y tener su información recopilada por teléfono. Las direcciones de estas operaciones que no parearon con las que ya estaban listadas en el Archivo Maestro Decenal de Direcciones fueron visitadas en una operación de Verificación en el Campo para determinar si existían. Las direcciones verificadas fueron añadidas al listado de direcciones.

Una fuente adicional de información sobre unidades de vivienda listadas en el Archivo Maestro Decenal de Direcciones es la operación Seguimiento por Falta de Respuesta (NRFU, por sus siglas en inglés). Durante NRFU, los enumeradores dieron seguimiento a unidades que no habían devuelto un cuestionario predirigido del censo. Las unidades en NRFU podrían ser eliminadas o consideradas como desocupadas. A la vez, las unidades que no aparecieron en el listado de direcciones o los mapas pudieron ser añadidas y enumeradas simultáneamente. Esta operación se llevó a cabo en las áreas de actualización/entrega.

### **ENUMERACIÓN DE LUGARES QUE PROVEEN SERVICIOS**

La Enumeración de Lugares que Proveen Servicios (SBE, por sus siglas en inglés) se diseñó para contar a las personas sin una residencia habitual que usan facilidades de servicios (por ejemplo, refugios, comedores de beneficencia y servicios rodantes de comida). Se enumeraron sólo las

---

personas que usaron la facilidad de servicios el día de la entrevista. Además, las personas enumeradas en localizaciones específicas al aire libre y las personas sin una residencia habitual que completaron cuestionarios de Sea Usted Contado (BCF, por sus siglas en inglés) aumentaron las cifras de SBE. Se incluyó el total final en la población total. Este componente de la enumeración *no* debe interpretarse como un recuento completo de la población sin hogar.

## **CONFIDENCIALIDAD DE LOS DATOS**

El Negociado del Censo ha modificado algunos datos en esta publicación de datos para proteger la confidencialidad. La Sección 9 del título 13 del Código de los Estados Unidos prohíbe que el Negociado del Censo haga público los resultados que puedan identificar datos sobre un individuo.

La comisión interna de revisión de supresión de datos (Internal Disclosure Review Board) del Negociado del Censo establece las reglas de confidencialidad para todas las emisiones de datos. El método de usar listas de cotejo se utiliza para asegurar que se consideran y se manejan todos los riesgos posibles a los cuales se expone la confidencialidad de los datos. Se pueden enviar preguntas sobre la confidencialidad a: [webmaster@census.gov](mailto:webmaster@census.gov) Attention Policy.

### **Título 13 del Código de los Estados Unidos**

El título 13 del Código de los Estados Unidos autoriza que el Negociado del Censo lleve a cabo censos y encuestas. La Sección 9 del mismo título requiere que cualquier información que se recopile del público bajo la autoridad del título 13 se mantenga confidencial. La Sección 214 del título 13 y las Secciones 3559 y 3571 del título 18 del Código de los Estados Unidos estipula que se impongan penalidades de hasta 5 años de cárcel y hasta \$250,000 en multas por la divulgación intencional de información confidencial del censo.

### **Límite a la divulgación de datos**

El límite a la divulgación de datos es el proceso usado para proteger la confidencialidad de los datos. Una divulgación de datos ocurre cuando alguien puede usar información estadística publicada para identificar a un individuo que ha provisto información bajo una promesa de confidencialidad. Utilizando los procedimientos de la limitación a la divulgación, el Negociado del Censo modifica o elimina las características que podrían arriesgar la confidencialidad si se revelan. Aunque pueda parecer que una tabla muestra información sobre un individuo específico, el Negociado del Censo ha tomado pasos para disfrazar los datos originales a la vez que asegura que los resultados todavía sean útiles.

### **Intercambio de datos**

El intercambio de datos es un método de limitar la divulgación de datos que se diseñó para proteger la confidencialidad en tablas con datos de frecuencia (el número o porcentaje de la población con ciertas características). El intercambio de datos se hace editando las fuentes de datos o intercambiando los registros por una muestra de casos al crear una tabla. Una muestra de hogares se selecciona y se para usando un conjunto seleccionado de variables claves con hogares en áreas geográficas vecinas que tienen características semejantes (como por ejemplo, el mismo número de adultos y el mismo número de hijos). Como el intercambio ocurre frecuentemente dentro de un área vecina, no afecta los totales marginales para el área o los totales que incluyen datos de múltiples áreas. Debido al intercambio de datos, los usuarios no deben suponer que las tablas con celdas con un valor de 'uno' o 'dos' revelan información sobre individuos específicos.

## **ERRORES AJENOS AL MUESTREO**

En cualquier operación estadística de gran escala, como el Censo 2000, ocurren errores humanos y de computadora. Estos errores frecuentemente se conocen como errores ajenos al muestreo. Tales errores incluyen el no enumerar cada hogar o a cada persona en la población, no obtener toda la información requerida de los respondedores, obtener información incorrecta o

---

imprecisa y anotar información incorrectamente. Además, pueden ocurrir errores durante la revisión en el campo del trabajo de los enumeradores, durante el manejo de los cuestionarios en la oficina o durante el procesamiento electrónico de los cuestionarios.

Como es imposible eliminar completamente los errores ajenos al muestreo de una operación tan grande y compleja como el censo decenal, el Negociado del Censo intenta controlar las fuentes de tales errores durante las operaciones de recopilación y procesamiento. A continuación, se describen las fuentes principales de errores ajenos al muestreo y los programas establecidos para controlar este tipo de error durante el Censo 2000. Sin embargo, el éxito de estos programas dependió de cuán bien se siguieron las instrucciones en la realidad durante el censo. Como parte del programa de evaluación del Censo 2000, se evaluarán tanto los efectos de estos programas como la cantidad de errores que permanecieron después de su aplicación.

### **Tipos de Errores Ajenos al Muestreo**

**Falta de Respuesta.** La falta de respuesta a ciertas preguntas en el cuestionario del censo o no obtener cualquier información sobre una unidad de vivienda da base a que se introduzca sesgo en los datos, ya que no se han observado las características de las personas que no han respondido y pueden diferir de aquéllas provistas por los respondedores. Como resultado, cualquier procedimiento de imputación utilizando los datos del respondedor puede que no refleje estas diferencias ya sea al nivel elemental (persona o unidad de vivienda individual) o como promedio. El minimizar la falta de respuesta ofrece alguna protección contra la introducción de grandes sesgos en los datos. Las características para las preguntas sin respuestas fueron imputadas utilizando datos reportados para una persona o unidad de vivienda con características similares.

**Errores del Respondedor y del Enumerador.** La persona que contestó el cuestionario y lo devolvió por correo o respondió a las preguntas hechas por un enumerador para un hogar podría ser una fuente de error. Aunque la manera en que se frasearon las preguntas fue probada en varios estudios experimentales antes del censo, el respondedor que contesta por correo puede dejar pasar o malinterpretar una pregunta, o contestar una pregunta en una manera que el sistema de la captura de datos no pueda interpretar correctamente. El enumerador también puede malinterpretar o anotar incorrectamente la información provista por el respondedor; o puede que no recopile alguna de la información para una persona u hogar, o puede recopilar datos para hogares que no habían sido designados como parte de la muestra. Para controlar problemas como estos durante la enumeración en el campo, se verificó el trabajo de los enumeradores cuidadosamente. El personal de campo se preparó para sus tareas utilizando adiestramientos estándares que incluían experiencia práctica en el uso de los materiales del censo. Una muestra de los hogares entrevistados por enumeradores fue entrevistada nuevamente para controlar la posibilidad de que los enumeradores estuvieran sometiendo datos falsificados.

**Errores de Procesamiento.** Las muchas etapas en el procesamiento de los datos censales representan fuentes potenciales para la introducción de errores ajenos al muestreo. El procesamiento de los cuestionarios censales completados por los enumeradores incluyó la revisión en el campo por el líder de grupo, la verificación de recibo de los cuestionarios y el envío de cuestionarios completados. No hubo revisiones en el campo para los cuestionarios devueltos por correo en este censo. También se puede introducir errores por la malinterpretación de datos por el sistema de captura de datos o por no capturar toda la información que los respondedores o enumeradores proveyeron en las formas. Las respuestas escritas fueron procesadas en operaciones de codificación, que también pueden ser una fuente de errores de procesamiento en los datos. Muchas de las operaciones de campo, codificación y computadora pasan por un número de revisiones para el control de la calidad para asegurar que fueron aplicadas correctamente.

### **Reducción de Errores Ajenos al Muestreo**

Para reducir los varios tipos de errores ajenos al muestreo, se utilizaron un número de técnicas durante la planificación, el desarrollo del listado de direcciones, la recopilación de datos y las actividades de procesamiento de datos. Se usaron métodos de verificar la calidad durante las

---

fases de recopilación y procesamiento de datos para mejorar la calidad de los datos. Se estableció un programa de entrevistas suplementarias para minimizar los errores en la fase de recopilación de datos para los cuestionarios completados por enumeradores.

Se establecieron varios programas para mejorar la cobertura durante el desarrollo del listado de direcciones y la enumeración y procesamiento del censo para minimizar la subcobertura de la población y unidades de vivienda. Estos programas se desarrollaron a base de experiencias del censo de 1990 y los resultados de la prueba del Censo 2000.

- Cuestionarios de Sea Usted Contado, formas sin direcciones que pidieron partidas de la forma corta del cuestionario, además de unas partidas adicionales, estaban disponibles en lugares públicos para personas que creían que no fueron contadas en el censo.
- Se entregó una carta de aviso a todas las direcciones en las áreas de actualización/entrega antes de que se entregaran los cuestionarios del censo. También se envió una tarjeta postal recordatoria a estas direcciones.
- Se enviaron por correo cuestionarios en inglés a aquellas personas que los pidieron.
- Se le dió mucha publicidad a un número de teléfono libre de cargos que estaba disponible para contestar preguntas sobre los cuestionarios. Además, se podían aceptar contestaciones por teléfono de hogares que habían recibido la forma corta del cuestionario.
- En el programa de Actualización Local de Direcciones Censales (LUCA, por sus siglas en inglés), muchos gobiernos locales tuvieron la oportunidad de expresar sus preocupaciones específicas sobre la precisión e integridad del Archivo Maestro de Direcciones antes de que entregaran los cuestionarios.

### **Resolución al Recibo de Múltiples Respuestas**

Con las múltiples maneras en que personas podían empezar su enumeración y las operaciones de seguimiento en el campo, fue muy probable que algunas personas fueran enumeradas más de una vez. Se estableció un proceso especial por computadora para controlar la extensión de estos tipos de errores ajenos al muestreo al resolver situaciones en las cuales se recibió más de un cuestionario de una dirección. El proceso consistió en varias etapas. Las direcciones que devolvieron más de una forma fueron analizadas. Los datos sobre hogares de un cuestionario fueron escogidos como los datos sobre hogares que se utilizarían en procesamientos subsiguientes del censo. Dentro de cada una de estas direcciones, se compararon los registros de persona en cada cuestionario devuelto con los registros de persona en los otros cuestionarios devueltos para la misma dirección. Las personas que se incluyeron en dos o más cuestionarios devueltos fueron marcadas como tal y sólo uno de los registros de persona se utilizó en procesamientos subsiguientes.

### **Imputación de la Clasificación de Unidades de Vivienda y Recuentos de Población**

Después de completar todas las actividades de recopilación de datos para el Censo 2000, se preparó un archivo de computadora de las unidades de vivienda del censo. Para algunas unidades de vivienda, no había información disponible sobre la clasificación de la unidad de vivienda, si estaba ocupada, desocupada o no existía. Se definieron estas unidades de vivienda como no clasificadas. Las unidades de vivienda no clasificadas fueron asignadas una clasificación de unidad de vivienda—ocupada, desocupada o no existente—asignándoles la clasificación de una unidad de vivienda cercana a la unidad no clasificada. Además, no se sabía el número de personas que vivían en algunas unidades de vivienda ocupadas. Se asignó el recuento de población de una unidad de vivienda ocupada cercana a las unidades de vivienda con una población desconocida. Todos los otros datos para estas unidades de vivienda fueron asignados por medio de sustitución o asignación durante la revisión de datos inaceptables que se explica en la sección a continuación.

### **REVISIÓN DE DATOS INACEPTABLES**

El objetivo de la operación de procesamiento fue producir un conjunto de datos que describe la población lo más claro y preciso posible. En un cambio importante de prácticas anteriores, la información en los cuestionarios del Censo 2000 generalmente no se revisó durante la

---

recopilación de datos en el campo ni durante operaciones de captura de datos para asegurar su uniformidad, integridad y aceptabilidad. Líderes de grupo del censo y oficinistas locales revisaron los cuestionarios completados por enumeradores para verificar que se siguieron procedimientos específicos. No se hicieron revisiones en la oficina de los cuestionarios devueltos por correo para asegurar que la información en el cuestionario se podría capturar. Además, no se comunicaron con hogares para recopilar datos que faltaron de los cuestionarios devueltos como se hizo en censos anteriores.

La mayoría de cuestionarios recibidos por correo de respondedores, al igual que los completados por enumeradores, fueron procesados por un nuevo sistema de escáner creado por contratistas que utilizó reconocimiento óptico de marcas y caracteres para convertir las respuestas en archivos de computadora. El reconocimiento óptico de caracteres, o OCR, utilizó varias verificaciones de diseño y contexto para estimar los límites de precisión para cada campo con respuestas escritas. El sistema también permitió soft edits en la mayoría de respuestas numéricas que fueron escritas por una persona que no fuera el respondedor para decidir si los valores de los campos leídos por la máquina eran aceptables. Si el valor leído tenía un límite de precisión más bajo que el límite aceptado o si estaba fuera de la gama de soft edit, se mostraba la imagen de la partida a un oficinista que entonces entraba la respuesta.

Para controlar la creación de personas posiblemente erróneas de cuestionarios completados incorrectamente o que contenían marcas fuera de los encasillados, se hizo una revisión del número de personas indicadas en cada forma devuelta por correo y cada cuestionario completado por un enumerador como parte del sistema de captura de datos. Si no pasaban esta revisión, un operador revisaba la imagen del cuestionario en su estación de trabajo e identificaba registros de personas erróneas y corregía errores de interpretación del reconocimiento óptico de caracteres en el campo de recuento de población.

En la oficina central del Negociado del Censo, los registros de datos sobre envíos por correo fueron sujetos a una revisión por computadora que identificó hogares que indicaron un problema posible de cobertura y aquéllos con más de seis miembros en el hogar, el número máximo de personas que podían ser enumeradas en un cuestionario enviado por correo. Se trató de comunicarse con estos hogares por teléfono para corregir la falta de uniformidad del recuento y para recopilar los datos censales de aquellas personas para quienes no había espacio en el cuestionario.

Para información incompleta o no uniforme en los registros de datos de cuestionarios se asignó valores aceptables utilizando procedimientos de imputación durante la revisión automatizada final de los datos recopilados. Con frecuencia, se necesitan imputaciones, o asignaciones computarizadas de códigos aceptables en lugar de anotaciones inaceptables o en blanco, cuando falta una anotación para una partida dada o cuando la información reportada para una persona en esa partida no concuerda con otra información para esa misma persona. Este proceso se conoce como asignación. Al igual que en censos anteriores, el procedimiento general para cambiar anotaciones inaceptables fue asignar una anotación para una persona que concordara con anotaciones para personas con características similares. La asignación de códigos aceptables en lugar de anotaciones en blanco o inaceptables aumenta la utilidad de los datos. Las tasas de asignación para partidas censales están disponibles con los datos censales publicados.

Otra manera en que se hicieron correcciones durante el proceso de revisión automatizada fue a través de la sustitución; es decir, la asignación de un conjunto completo de características para las personas en un hogar. Cuando había indicación de que una unidad de vivienda estaba ocupada por un número especificado de personas pero el cuestionario no contenía la información para las personas dentro del hogar o cuando los ocupantes no estaban listados en el cuestionario, se seleccionó como sustituto un hogar del mismo tamaño y se duplicó el conjunto completo de características del sustituto. No se sustituyen características de vivienda. Matriz H18, Unidades de Vivienda Ocupadas Sustituídas, representa el recuento de unidades de vivienda en las cuales se sustituyeron las personas.